Structure and Expression of the Vaccinia Virus Gene Which Prevents Virus-Induced Breakdown of RNA

RADIYA F. PACHA,[†] RONALD J. MEIS, AND RICHARD C. CONDIT^{‡*}

Department of Biochemistry, State University of New York at Buffalo, 140 Farber Hall, Buffalo, New York 14214

Received 23 February 1990/Accepted 11 May 1990

Three noncomplementing vaccinia virus temperature-sensitive mutants, *ts4*, *ts22*, and *ts23*, exhibit an abortive late phenotype characterized by the simultaneous cessation of protein synthesis, the breakdown of rRNA and viral mRNA, and an increase in intracellular concentrations of 2'-5'-linked oligoadenylates late during infection at the nonpermissive temperature (R. F. Pacha and R. C. Condit, *J. Virol.* 56:395–403, 1985; R. J. Cohrs, R. C. Condit, R. F. Pacha, C. L. Thompson, and O. K. Sharma, *J. Virol.* 63:948–951, 1989). We have identified the virus gene affected by the abortive late mutants, determined its DNA sequence, and analyzed its transcription. The gene resides in the *Hind*III A DNA fragment, it has a predicted coding capacity of 57 kilodaltons, and it is transcribed both early and late during infection. The early transcript of the abortive late gene is unusual; it contains a 426-nucleotide 5' untranslated region, and it must be synthesized by transcription through an early transcription termination signal which is located in the middle of the gene in a hairpin loop structure. DNA sequence and transcription analysis of two flanking genes is also presented.

Vaccinia virus, the laboratory prototype of poxviruses, replicates its 186-kilobase-pair (kb) DNA genome in the cytoplasm of susceptible host cells (reviewed in reference 23). While replication of a DNA genome in a cytoplasmic compartment has resulted in the evolution by the virus of a high degree of independence from host nuclear DNA and RNA metabolism enzymes, infection by vaccinia virus nevertheless represents a complex interaction between viral and host functions. A striking example of the virus-host interaction is revealed through biochemical analysis of infection by vaccinia virus temperature-sensitive (*ts*) mutants bearing an abortive late phenotype.

The abortive late phenotype is displayed during infection under nonpermissive conditions by each of three vaccinia *ts* mutants, *ts4*, *ts22*, and *ts23*, which together comprise a single complementation group (5, 6, 24; R. C. Condit, unpublished). Infection by abortive late mutants of vaccinia virus under nonpermissive conditions starts normally: host protein synthesis is shut off, early viral protein synthesis starts and stops normally, and late viral protein synthesis is initiated at the proper time during infection. The hallmark of the abortive late phenotype is the breakdown of viral mRNA and host rRNA late during infection, accompanied by a cessation of all protein synthesis in infected cells. A closely related phenotype can be induced during infection by wild-type (wt) vaccinia virus if infections are done in the presence of the anti-poxvirus drug isatin- β -thiosemicarbazone (IBT).

Recent evidence suggests that the RNA breakdown observed during an abortive late infection is catalyzed by a host enzyme, the 2'-5'-oligoadenylate synthase-dependent RNase L (4). Investigation of this possibility was prompted by the fact that the rRNA breakdown observed during infection by abortive late mutants of vaccinia virus is virtually identical to what is observed when vaccinia virus infection is inhibited

by interferon treatment of mouse L cells (11). Interferon-induced rRNA breakdown in virus-infected cells is thought to result from a cascade of events which include induction of a double stranded RNA-activated 2'-5'-oligoadenylate synthase, synthesis from ATP of 2'-5'-linked adenylate oligomers (2-5A), and activation of RNase L by 2-5A (reviewed in reference 18). Cohrs et al. (4) have shown that intracellular concentrations of 2-5A are abnormally high during infection by abortive late mutants of vaccinia virus under conditions during which rRNA breakdown is observed. Not only do these results implicate the 2-5A cascade in the abortive late phenotype, but they suggest that an interaction between viral gene products and the host 2-5A–RNase L pathway may be an important feature of a normal vaccinia virus infection. The promise that an understanding of the viral gene(s) involved in the abortive late phenotype may provide insight into this novel virus-host interaction has prompted us to characterize the gene affected by the abortive late mutants.

We report here the mapping, DNA sequence, and transcription analysis of the vaccinia virus gene affected by the abortive late mutants. The DNA sequence and transcription analysis of two flanking viral genes is also presented.

MATERIALS AND METHODS

Cells and virus. BSC40 cells, wt vaccinia virus strain WR, *ts22*, *ts23*, and the conditions for their growth, infections, and plaque titrations have been previously described in detail (5, 6). The spontaneous IBT-resistant vaccinia virus mutant, IBT^{rr90} (a gift of Zahra Fathi), was isolated from plaque titrations of wt vaccinia virus done on BSC40 cells at 31°C in the presence of 90 μ M IBT. The mutant was repeatedly plaque purified at 31°C in the presence of 90 μ M IBT and was grown at 31°C in the presence of 90 μ M IBT.

Marker rescue. Marker rescue was done as previously described (35). Briefly, cells were infected with *ts* virus at an appropriate multiplicity of infection determined empirically by terminal dilution, incubated at the permissive temperature for 2 to 6 h, transfected with calcium phosphate-precipitated supercoiled plasmid or single-stranded bacteriophage DNA cloned from wt virus, and incubated at the

* Corresponding author.

[†] Present address: Institute of Molecular Virology, University of Wisconsin, Madison, WI 53706.

[‡] Present address: Department of Immunology and Medical Microbiology, Box J-266 JHMC, University of Florida, Gainesville, FL 32610.

nonpermissive temperature for 4 days. The dishes were then stained with crystal violet, and plaques were counted. Incubation at the permissive temperature for up to 6 h before transfection and shift up seemed to be important for obtaining a strong positive marker rescue of *ts23*.

Nucleic acid purification. Vaccinia virus genomic DNA and recombinant DNA clones in cosmid vectors were purified as previously described (35). Single-stranded and replicative-form M13 DNA and bacterial plasmids were grown and purified by standard procedures (1). Total RNA was purified from vaccinia virus-infected (multiplicity of infection, 15 PFU per cell) or mock-infected cells as described previously (24). Incubations were at 37°C. Early viral RNA samples were purified from cells infected under three different conditions. Authentic early viral RNA samples were isolated at 2 h postinfection in the absence of any drug, cycloheximide early viral RNA was isolated from cells infected for 8 h in the presence of 100 µg of cycloheximide per ml, and 1-β-D-arabinofuranosylcytosine (araC) early RNA was isolated from cells infected for 7 h in the presence of 60 µg of araC per ml. Late viral RNA was isolated from cells infected for 8 h in the absence of any drug, and mock RNA was isolated at 4 h after mock infection of cells in the absence of drug. Poly(A)⁺ RNA was isolated by oligo(dT)-cellulose chromatography as described in the Bethesda Research Laboratories, Inc., product profile for oligo(dT)-cellulose.

Recombinant DNA clones for marker rescue analysis. The map positions of recombinant DNA clones used for marker rescue and DNA sequencing are shown in Fig. 1 and 3A. Viral DNA for the clone pB¹ 14.0 was obtained from the vaccinia virus mutant IBT^{TP90} with the intention of testing not only for rescue of *ts22* but also for IBT resistance (data not shown). DNA for all other clones was obtained from wt virus. Clones were constructed by using standard protocols (1, 22).

The clones pWR 93-130, pWR 112-149, and pWR 120-150 have been described previously (35). pB¹ 14.0 contains a 14.0-kb *Bam*HI subfragment of the vaccinia virus *Hind*III A fragment, cloned from IBT^{TP90} viral DNA into the *Bam*HI site of pUC13. pBS 6.0 contains a 6.0-kb *Bam*HI-*Sph*I fragment from pWR 112-149 DNA cloned between the *Bam*HI and *Sph*I sites of pUC18. pBP 3.8, which contains a 3.8-kb *Bam*HI-*Pvu*II fragment cloned into pBR322 (36), was obtained from R. Wittek (Universite de Lausanne, Lausanne, Switzerland). pX 5.0 contains a 5.0-kb *Xba*I fragment from pWR 93-130 DNA cloned into the *Xba*I site of pUC18. mXS 3.0 contains a 3.0-kb *Xba*I-*Sph*I fragment from pBS 6.0 DNA cloned between the *Xba*I and *Sph*I sites of M13 mp19. mSX 3.0 contains the end-filled 3.0-kb *Xba*I-*Sph*I fragment from mXS 3.0, cloned in the opposite orientation relative to mXS 3.0 in the *Sma*I site of the vector M13 mp18. mR₁ 1.4 and mR₂ 1.4 contain the 1.4-kb *Eco*RI fragment from pX 5.0 cloned in both orientations into the *Eco*RI site of M13 mp18. φ74 and φ349 are exonuclease deletion subclones of mXS 3.0 and mSX 3.0, respectively, generated during DNA sequence analysis of the 3.0-kb *Xba*I-*Sph*I fragment. φ349 contains DNA from the *Xba*I site at position 1 in the sequence shown in Fig. 2 through position 1463, 9 nucleotides downstream of the open reading frame (orf) Ac R AUG. φ74 contains DNA from position 1454, 6 nucleotides downstream of the orf Ac R AUG through the *Sph*I site at position 2955.

DNA sequence analysis. Replicative-form DNAs from the clones mXS 3.0 or mSX 3.0 were linearized with *Bam*HI and *Kpn*I or *Sal*I and *Pst*I, respectively, and treated with exonuclease III to produce overlapping deletion subclones for DNA sequence analysis (12). DNA sequence analysis was

done by using the chain termination method of Sanger et al. (30). A region of approximately 30 nucleotides which contains a hairpin loop structure (see Fig. 3B) was refractory to dideoxy sequencing. Therefore, an exonuclease III deletion clone spanning this region was sequenced, using the chemical method of Maxam and Gilbert (21). The DNA sequence reported in Fig. 2 has been submitted to GenBank and assigned the accession number M32064.

For transcription mapping experiments, DNA sequence analysis was done as described in the Sequenase manual (United States Biochemical Corp.), using appropriate combinations of oligonucleotide primers and DNA templates.

Oligonucleotides. Oligonucleotides P1, P2, P3, and P4 were synthesized on a Biosearch 8600 oligonucleotide synthesizer and purified by gel electrophoresis as described in the Biosearch 8600 instruction manual. The positions of the oligonucleotides relative to the sequenced *Xba*I-*Sph*I fragment are shown in Fig. 2.

Northern (RNA) analysis. Appropriate subclones of mXS 3.0 were constructed in the vector pT7/T3 (Bethesda Research Laboratories). Continuously labeled RNA probes were synthesized as described in the Promega Biotec catalog, transcription systems applications protocol no. 1, using T7 or T3 RNA polymerase (New England BioLabs); DNA templates were linearized with appropriate restriction endonucleases. Coordinates and orientations of the probes are detailed in Fig. 4E.

RNA samples were treated with formaldehyde and electrophoresed essentially as described in unit 4.9 of Ausubel et al. (1), except that electrophoresis was done on 1.5% agarose gels in 15 mM phosphate buffer. Then, 5.6 µg of total RNA or 1 to 3 µg of poly(A)⁺ RNA was run per lane. RNA was transferred to hybridization membranes (GeneScreen; New England Nuclear) by using the capillary blot procedure, and blots were prehybridized, washed, and autoradiographed using method III described in the GeneScreen hybridization membrane instruction manual. Radiolabeled probe (10⁷ cpm) was used for each hybridization reaction.

S1 nuclease and primer extension analysis. 5' End-labeled, single-stranded DNA probes for nuclease S1 analysis were synthesized by elongating end-labeled oligonucleotide primers annealed to single-stranded M13 DNA templates, followed by restriction endonuclease cleavage and gel purification essentially as described in unit 4.6 of Ausubel et al. (1). Asymmetrically 5' or 3' end-labeled double-stranded DNA fragments were prepared by standard techniques (1). Precise coordinates of the S1 probes are provided in Fig. 5 and 6. S1 analysis was done as described in unit 4.6 of Ausubel et al. (1). Twenty-five micrograms of total RNA or 3 to 5 µg of poly(A)⁺ RNA was used in each reaction. Hybridization conditions are detailed in the figure legends. Nuclease-protected products were electrophoresed on an 8% denaturing polyacrylamide sequencing gel, fixed for 30 min in 5% methanol-5% acetic acid, dried, and autoradiographed.

Primer extension analysis was done essentially as described in unit 4.8 of Ausubel et al. (1). Extended primers were analyzed by electrophoresis as described under S1 analysis.

Computer analysis. Computer analysis was done on a DEC VAX 8650, using software obtained from Roger Staden (34), the University of Wisconsin Genetics Computer Group (9), and the National Biomedical Research Foundation.

RESULTS

Localization and DNA sequence analysis of the abortive late gene. Coarse marker rescue mapping done with wt vaccinia

FIG. 1. Map positions of subclones from the *Hind*III A fragment. From the top down are shown: map positions of three cosmid clones; a *Hind*III map of the vaccinia virus genome, including *Bam*HI sites within the *Hind*III A fragment (●); an expanded restriction map of the 14-kb DNA fragment contained in pB¹ 14.0; and map positions of five subclones spanning the left-hand one-half of pB¹ 14.0. B, *Bam*HI; X, *Xba*I; P, *Pvu*II; S, *Sph*I; E, *Eco*RI. X* indicates that we are uncertain whether this *Xba*I site is located as indicated or 700 nucleotides to the right of this position.

virus DNA clones in cosmid vectors previously positioned the *ts22* mutation within the left-hand one-third of the *Hind*III A fragment (35) (Fig. 1). Increased resolution of the map position of *ts22* was obtained by subcloning fragments from the *Hind*III A fragment and testing these subclones for the ability to rescue *ts22* (Fig. 1, Table 1). Positive marker rescue results were obtained with five overlapping subclones of the *Hind*III A fragment, represented by clones pB¹ 14.0, pBS 6.0, pX 5.0, mXS 3.0, mR₁ 1.4, and mR₂ 1.4. The map position of *ts22* is resolved by these experiments to a 1.4-kb *Eco*RI fragment represented by mR 1.4.

To determine the genetic organization of the region of the vaccinia virus genome which encodes *ts22*, the vaccinia virus DNA contained in clone mXS 3.0 was sequenced (Fig. 2 and 3). Computer analysis of the sequence revealed three orfs, designated from left to right as orf Aa L, orf Ab L, and orf Ac R. orf Aa L is translated in a leftward direction and is incomplete. orf Ab L is translated in a leftward direction and could encode a protein of 23-kilodaltons (kDa) molecular size. orf Ac R is translated in a rightward direction and could encode a protein of 57-kDa molecular size. Two additional features of the sequence are noteworthy when considered in the context of the hypothetical translation products. The first is the sequence TAAATG (the complement is shown in Fig. 2), from positions 1442 to 1437, in which the ATG is the hypothetical initiation codon for translation of orf Ab L. This sequence has been found in association with many but not all early and late vaccinia virus transcription promoters (13, 16, 17, 29). The association with late promoters seems to be more frequent than that with early promoters. Evidence has been presented which suggests that the sequence is necessary but not sufficient for the function of at least one late promoter and that the sequence may play a role in the synthesis of the 5' poly(A) head found on several late vaccinia virus mRNAs (2, 8, 13a, 32). The second sequence of note is an early viral transcription termination signal (28,

37), TTTTTTT, located at positions 2254 to 2260, in the middle of orf Ac R. The sequence is an integral part of a 15-base-pair stem and loop structure (Fig. 3B), which could influence its function during transcription of orf Ac R.

Comparison of the predicted protein sequences with sequences in the protein sequence data base from the Protein Identification Resource of the National Biomedical Research Foundation (release 22.0 of "protein" and release 40.0 of "new"), using the program FastA (9, 19), revealed no significant homologies.

To determine which orf is responsible for the abortive late phenotype, marker rescue of *ts22* and *ts23* was done with two orf-specific M13 DNA clones, ϕ 349, which corresponds to orfs Aa L and Ab L, and ϕ 74, which corresponds to orf Ac R (Fig. 3C, Table 1). The results show clearly that both *ts22*

TABLE 1. Marker rescue mapping of *ts22*^a

Clone	Plaques/dish
pWR 93-130	>200
pWR112-149	>200
pWR 120-150	0
pB ¹ 14.0	>200
pBS 6.0	100 ^b
pBP 3.8	0
mR ₁ 1.4	22 ^b
mR ₂ 1.4	12 ^b
pX 5.0	100 ^b
mXS 3.0	200 ^b
mSX 3.0	80 ^b
ϕ 349	0
ϕ 74	100 ^b

^a Marker rescue experiments were done as described in Materials and Methods, and the number of plaques appearing on each stained dish is tabulated. Physical map positions of the plasmid and phage DNAs are noted in Fig. 1 and 3.

^b Plaques on these dishes were uniformly smaller than wt plaques.

FIG. 2. Nucleotide sequence of the 3-kb *Xba*I-*Sph*I DNA fragment containing the abortive late gene. The nucleotide sequence of the insert in mXS 3.0 and its most probable translation are shown. The initiation codons for three orfs (Aa L, Ab L, and Ac R) and their directions of translation are indicated below the amino acid sequence. The orf Ac R early mRNA 5' end is indicated with asterisks above the nucleotide sequence at positions 1027 and 1029. The positions of four oligonucleotide primers (P1, P2, P3, P4) used for transcription mapping are shown above the DNA sequence. An early transcription termination signal in orf Ac R at positions 2254 to 2260 is overlined.

and *ts23* were rescued with $\phi 74$ and not with $\phi 349$ and therefore that orf Ac R is responsible for the abortive late phenotype.

Attempts to rescue all three abortive late mutants, *ts4*, *ts22*, and *ts23*, with cosmid clones or with the *Hind*III A fragment subclones revealed potentially important genetic characteristics of each mutant. First, we noted that rescue of

ts22 with clones smaller than pB^I 14.0 consistently yielded plaques which were smaller than wt plaques grown under the same conditions (Fig. 3C). We passaged virus from both large plaques obtained from rescue with pWR 93-130 and small plaques obtained from rescue with $\phi 74$, and we found that the large- and small-plaque morphologies were stable characteristics of the rescued virus (data not shown). A

```

1981  TCCGGATGTATTAATAGTAGTCAGTAGACATCTGACAAACGATGCCTTTTGTAATATATCAATAAGCATTATGATTGTTCATCTTGA 2070
      P D V L I V V S R H L T N D A F C K Y I N K H Y D L F I L D -
2071  TGAATCACATACGTATAATCTGATGAACAATACAGCAGTTACAAGATTTTGTAGCTATTATCCTCCGATGATGTGTTATTTTAACTGC 2160
      E S H T Y N L M N N T A V T R F L A Y Y P P M M C Y F L T A -
2161  TACACCTAGACCAGCTAACCGAATTTATTGTAACAGTATTATTAATATTGCCAAGTTATCCGATCTAAAAAACTATCTATGCGGTAGA 2250
      T P R P A N R I Y C N S I I N I A K L S D L K K T I Y A V D -
      -----
2251  TAGTTTTTTTGAGCCATATTCCACAGACAATATTAGACATATGGTAAAACGACTAGTAGGACCATCTAATAAATATCATATATATACCGA 2340
      S F F E P Y S T D N I R H M V K R L D G P S N K Y H I Y T E -
2341  GAAGTTATTATCTGTAGACGAGCCTAGAAATCAACTTATCTTAATACCTGGTAGAAGAATTCAAGTCAGGAACCTATTAATCGCATTTT 2430
      K L L S V D E P R N Q L I L N T L V E E F K S G T I N R I L -
2431  AGTTATTACTAACTACGTGAACATATGGTATTATCTCAACAGGATTATAGATCTTTTCGGACCAGAGGTTGTATTATAGGAGACGC 2520
      V I T K L R E H M V L F Y K R L L D L F G P E V V F I G D A -
2521  CCAAAATAGACGTACTCCAGATATGGTCAAATCAATCAAGGAACCTAAATAGATTATATTCGTATCCACCTTATTTTATCCGGTACTGG 2610
      Q N R R T P D M V K S I K E L N R F I F V S T L F Y S G T G -
2611  TTTAGATATTCCTAGTTTGGATTGCTGTTTCATTGCTCGGCAGTAATCAACAATATGCAATAGACAAATTACTAGGAGGGTATGTGC 2700
      L D I P S L D S L F I C S A V I N N M Q I E Q L L G R V C R -
2701  AGAAACAGAACTATTAGATAGGACGGTATATGTATTCTTAACACATCCATCAAGAAATAAAGTACATGATAGGAAATTCATGCAACG 2790
      E T E L L D R T V Y V F P N T S I K E I K Y M I G N F M Q R -
2791  AATTATTAGTCTGTCTGTAGATAAACTAGGATTAAACAGAAAGTTATCGGAACATCAAGAATCCGATCCCACTTCTGTATGTACAAC 2880
      I I S L S V D K L G F K Q E S Y R K H Q E S D P T S V C T T -
2881  ATCCTCCAGAGAAGACGTGTATTAAATAGAATATTAACTCGCAAAATCGTTAAGAAGTTAAGCGACGATCCG 2955
      S S R E E R V L N R I F N S Q N R *

```

FIG. 2—Continued.

likely explanation for this finding is that *ts22* actually contains two mutations, a *ts* mutation in orf Ac R and another temperature-independent plaque morphology mutation which maps outside of the region defined by pBS 6.0 but within the region defined by pB^I 14.0. This possibility provided strong incentive to map an additional abortive late mutant to a single orf and thereby to confirm that orf Ac R was responsible for the abortive late phenotype. Marker rescue mapping of *ts4* and *ts23* was complicated by the fact that both of these mutants exhibit consistently high reversion indices ($>10^{-5}$) and that the revertants contribute to a high background in the marker rescue assay. The background in experiments with *ts4* was never low enough to yield a convincing positive result. The background (Fig. 3C, -DNA) in rescue experiments involving *ts23* was consistently high, but it was low enough that convincing positive rescue results were obtained. The results with *ts23* showed not only that the mutant is rescued with $\phi 74$ but also that the resulting plaques have a wt plaque morphology, confirming that orf Ac R is responsible for the abortive late phenotype.

Northern analysis. Transcriptional organization of the sequenced 3.0-kb *XbaI-SphI* DNA fragment was first assessed by Northern blot analysis, using single-stranded riboprobes specific for each of the hypothetical orfs (Fig. 4). Use of a riboprobe designed to detect rightward transcripts through orf Ac R revealed that orf Ac R is transcribed both early and late during vaccinia virus infection (Fig. 4A). A major early transcript of approximately 2,400 nucleotides in length was detected in total RNA extracted from cells between 1 and 2 h after infection. The same signal was observed in RNA extracted from infected cells treated with cycloheximide or araC, treatments which permit early but not late viral gene expression. Late during infection a heterogeneous smear of RNA from orf Ac R was detected, ranging in size from approximately 1,500 to 5,000 nucleotides. The observed heterogeneity of RNA size is a well-documented characteristic of late vaccinia virus mRNA, resulting from the absence of discrete termination of transcription late during viral

infection (7, 20). Northern analysis using a riboprobe designed to detect RNA that would be transcribed through orf Ac R in a leftward direction revealed only late transcripts ranging in size from approximately 1 to 4 kb (Fig. 4B). This suggests that a late gene is located downstream from orf Ac R and that transcripts from this gene extend into the orf Ac R coding region. Northern blots using riboprobes designed to detect transcripts from orfs Aa L and Ab L revealed only late transcripts from each of these orfs, ranging in size from approximately 1 to 4 kb (Fig. 4C and D). In summary, the Northern analysis indicated that orfs Aa L and Ab L are late viral genes, whereas orf Ac R is expressed both early and late during viral infection.

S1 and primer extension analyses. (i) **Early RNA.** The 5' end of the orf Ac R early transcript was mapped by both primer extension and S1 analysis. Preliminary experiments using S1 probes or labeled primers with 5' ends located within the coding region of orf Ac R failed to detect an AUG-proximal early 5' end but showed instead a 5' end located approximately 400 nucleotides upstream from the orf Ac R AUG. Fine mapping of this early mRNA 5' end is presented in Fig. 5A. Primer extension analysis using a 5' end-labeled primer, P4 (Fig. 2), produced two bands which, when aligned with DNA sequence obtained with primer P4, correspond to positions 1027 and 1029 of the sequence shown in Fig. 2. This result is confirmed by S1 analysis using a single-stranded DNA probe extended from the same end-labeled primer, P4. These results position an early mRNA 5' end at nucleotides 1027 and 1029, 426 nucleotides upstream from the orf Ac R AUG.

The orf Ac R early mRNA 3' end was mapped by S1 analysis (Fig. 5B). Use of a 2.5-kb double-stranded DNA fragment asymmetrically 3' end labeled at position 2482 of the sequence shown in Fig. 2 yielded a 555-nucleotide S1-resistant DNA fragment. This result positions an early mRNA 3' end approximately 80 nucleotides downstream from the right-hand end of the sequence shown in Fig. 2. S1 experiments using 3' end-labeled probes spanning the inter-

FIG. 3. Physical mapping of the abortive late gene. (A) Summary map of the 3.0-kb *XbaI-SphI* DNA fragment containing the abortive late gene. From top to bottom: a scale for the 3.0-kb *XbaI-SphI* fragment; a diagram of the positions of the orfs in the 3.0-kb *XbaI-SphI* fragment, with boxes above center representing leftward-reading orfs and the box below center representing a rightward-reading orf; map position of two sequencing phages, $\phi 74$ and $\phi 349$, used to map *ts22* and *ts23* by marker rescue; map positions of early and late transcripts. Arrowheads indicate 3' ends, open circles indicate ends mapped by S1 analysis, hatched circles indicate ends mapped by primer extension, dashed lines indicate uncertainty in RNA length. (B) A hypothetical hairpin structure formed within orf Ac R, using nucleotides 2225 to 2263. (C) Marker rescue of *ts22* and *ts23* with orf-specific DNA clones. Marker rescue was done as described in Materials and Methods. Crystal violet-stained dishes are shown. Each row of dishes was infected with the mutant indicated at the left of the figure. Each column of dishes was transfected with the DNA indicated at the bottom of the figure.

nal transcription termination signal at positions 2254 to 2260 failed to detect an mRNA 3' end which would correspond to termination near this sequence (data not shown).

The 5' and 3' end analyses of orf Ac R early RNA together predicted a single 2,000-nucleotide transcript containing a 426-nucleotide 5' untranslated region, the orf Ac R coding

region, and a 100-nucleotide 3' untranslated region. These results are reasonably consistent with the observation of a 2,400-nucleotide early orf Ac R RNA on Northern blot analysis. The apparently larger size of the transcript on Northern blots can be explained by the presence in the RNA of 3' poly(A) sequences which would be detected on Northern blots but not with S1 analysis and of experimental error in size measurements of the RNA on Northern blots.

(ii) **Late RNA.** We have attempted to map the 5' end of the orf Ac R late transcript by both S1 and primer extension analyses (Fig. 6A). S1 analysis was done by using a double-stranded DNA probe asymmetrically 5' end labeled at position 1514 relative to the sequence shown in Fig. 2. Protected fragments of approximately 68 nucleotides in length were observed which would correspond to an RNA 5' end which maps at position 1446 of the sequence in Fig. 2, 6 nucleotides upstream from the orf Ac R AUG codon. Primer extension analysis using primer P3 (Fig. 2) yielded different results, depending on whether total RNA or poly(A)⁺ RNA was used as a template. With total RNA, extended products of approximately 73 nucleotides in length were observed, which would correspond to an RNA 5' end located at position 1436 of the sequence in Fig. 2, 13 nucleotides upstream from the orf Ac R AUG codon. This signal was not observed when poly(A)⁺ RNA was used as a template for reverse transcription. Instead, longer products were observed which would correspond to RNA 5' ends located at positions 1090 and 1145 of the sequence shown in Fig. 2, approximately 350 nucleotides upstream from the orf Ac R AUG codon.

Attempts to define an orf Aa L late mRNA 5' end by S1 and primer extension analyses using primer P1-derived probes yielded ambiguous results (data not shown). Although primer extension experiments yielded some AUG-proximal signals, these signals were not confirmed by S1 analysis. Instead, S1 analysis yielded signals corresponding

FIG. 4. Northern analysis of transcripts from the *XbaI-SphI* DNA fragment. Total RNA from wt-infected or mock-infected cells was transferred to hybridization membranes and hybridized with a ^{32}P -labeled riboprobe. Autoradiograms of Northern blots are shown. (A) orf Ac R sense RNA (probe A). The blot on the left shows results obtained with RNAs synthesized under various conditions of infection. Lanes: M, RNA from mock-infected cells; Ea, araC early RNA; Ec, cycloheximide early RNA; E, authentic early RNA; L, late RNA. The autoradiogram on the right shows results obtained with total RNA extracted from cells at different times postinfection, indicated at the top of the autoradiogram in hours. (B) orf Ac R antisense RNA (probe B). (C) orf Ab L sense RNA (probe C). (D) orf Aa L sense RNA (probe D). The autoradiograms in panels B, C, and D show results obtained with total RNA extracted from cells at different times postinfection, indicated at the tops of the autoradiograms in hours. (E) A cartoon showing the map positions of

riboprobes relative to the sequence organization of the 3.0-kb *XbaI-SphI* DNA fragment. The arrowhead represents the 3' end of the riboprobe. The coordinates of the riboprobes relative to the sequence in Fig. 2 are: probes A and B, nucleotides 1976 to 2400; probe C, nucleotides 1030 to 1514; and probe D, nucleotides 1 to 691.

to a position approximately 175 base pairs upstream of the orf Aa L AUG.

S1 and primer extension experiments designed to detect the orf Ab L late mRNA 5' end are shown in Fig. 6B. Primer extension was done with primer P2 (Fig. 2), and S1 analysis was done with a probe extended from the same primer, P2. S1 analysis yielded a strong signal which would correspond to an RNA 5' end which maps immediately upstream of the Ab L AUG codon. Primer extension analysis yielded a series of bands which together form a ladder of approximately 35 nucleotides in length. The 3' proximal end of the ladder maps immediately upstream of the Ab L AUG codon. Longer

primer extension products observed with poly(A)⁺ RNA correspond to positions in the DNA sequence ladder with bands in each lane and therefore may represent RNA structure-related stops in reverse transcription of long RNAs.

The results of the transcription analysis are summarized in Fig. 3 and discussed in more detail below.

DISCUSSION

We have identified and sequenced the gene affected by vaccinia virus mutants which exhibit an abortive late phenotype. The sequence analysis predicts that the gene encodes a protein of 57-kDa molecular weight. The predicted translation product did not show significant homology to any sequences presently contained in the PIR data base, and there are no striking structural features of the predicted polypeptide which give insight into its biochemical role in a vaccinia virus infection. An understanding of the function of

FIG. 5. 5' and 3' end analysis of the orf Ac R early mRNA. (A) 5' End analysis. Primer extension was done using end-labeled primer P4 (Fig. 2). S1 analysis was done using a single-stranded probe extended from primer P4, corresponding to nucleotides 815 to 1170 of the sequence shown in Fig. 2. Hybridization for the S1 analysis was done in 80% formamide at 42°C overnight. Autoradiograms are shown with extended or protected fragments aligned with the DNA sequence shown between the autoradiograms. Asterisks indicate the map positions of the 5' ends. Lanes: Ec, cycloheximide early RNA; Ec⁺, poly(A)⁺ cycloheximide early RNA; G, A, T, C, dideoxy sequencing reactions. (B) S1 analysis of the orf Ac R early mRNA 3' end. The probe was a 2.5-kb double-stranded DNA fragment from pX 5.0, asymmetrically 3' end labeled with [α -³²P]TTP at a Bgl/II site at position 2482 (Fig. 5C). Hybridization was done in 70% formamide at 37°C overnight. An autoradiogram is shown with positions of labeled size markers indicated at the right. Lanes: E, authentic early RNA; Ec, cycloheximide early RNA; Ec⁺, poly(A)⁺ cycloheximide early RNA; E⁺, poly(A)⁺ authentic early RNA; t, tRNA. (C) A cartoon showing the positions of the primers, the S1 probes, and the extended and protected products relative to the sequence organization of the 3.0-kb XbaI-SphI DNA fragment.

the abortive late gene product during a normal vaccinia virus infection will require direct analysis of the protein product in infected cells and analysis of virus mutants which may be related to the abortive late mutants, such as mutants dependent on or resistant to IBT.

It is interesting that the abortive late gene is transcribed both early and late during infection, since the phenotype expressed by the mutants occurs well after late viral transcription has begun (24). We have previously hypothesized that expression of the abortive late phenotype may require at least two steps, namely, induction of the cascade which precipitates RNA breakdown and protection from breakdown afforded by the abortive late gene product. The timing of the observed RNA breakdown probably reflects the time course of induction; the necessity for synthesis of the

FIG. 6. S1 and primer extension of late RNA 5' ends. (A) orf Ac R. Primer extension was done using end-labeled primer P3 (Fig. 2). S1 analysis was done using a 484-base-pair double-stranded DNA probe asymmetrically 5' labeled with [γ - 32 P]ATP and polynucleotide kinase at a *HincII* site at position 1514 relative to the sequence shown in Fig. 2. Hybridization for the S1 analysis was done in 53% formamide at 65°C overnight. Autoradiograms are shown. For the two lanes labeled S1 nuclease, the mobilities of size markers, with sizes given in nucleotides, are shown. Lanes: —, no RNA added; L, total late RNA. For the eight lanes labeled primer extension, G, A, T, and C are dideoxy sequencing reactions; L⁺ is poly(A)⁺ late RNA, and L is total late RNA. The duplicate L and L⁺ lanes represent two experiments done with the same labeled primer. Numbers to the right of the autoradiogram show the map positions of the indicated bands relative to the sequence in Fig. 2. In the expanded sequence, the orf Ab L and orf Ac R initiation codons and directions of translation are shown and the position of the orf Ac R primer extension signal is indicated with asterisks. (B) orf Ab L. Primer extension was done using end-labeled primer P2 (Fig. 2). S1

analysis was done using a probe extended from primer P2, corresponding to nucleotides 1380 to 1667 of the sequence shown in Fig. 2. Hybridization for S1 analysis was done in 80% formamide at 42°C overnight. Autoradiograms are shown and are labeled as in panel A. For the six lanes labeled S1 nuclease, the duplicate L lanes represent different amounts of the same material loaded on the gel. In the expanded sequence, the orf Ab L and orf Ac R initiation codons and the directions of translation are shown. The arrow to the right of the primer extension experiment indicates the position of the S1 signal. (C) A cartoon showing the positions of the primers, the S1 probes, and the extended and protected products relative to the sequence organization of the 3.0-kb *Xba*I-*Sph*I DNA fragment.

abortive late gene product at earlier times during infection awaits a more detailed description of its role in a normal viral infection.

Termination of early transcription during vaccinia virus infection is mediated by the sequence UUUUUNU in the RNA transcript, found 30 to 50 nucleotides upstream from the 3' end of the transcript (28, 33, 37). The abortive late gene, orf Ac R, is transcribed early as well as late during infection and yet contains an early transcription termination sequence (TTTTTTT) in the middle of the coding sequence, 800 nucleotides downstream from the AUG. Neither S1 nor Northern analysis yields any evidence that this signal is active in vivo in terminating early transcription of orf Ac R. Interestingly, this internal early transcription termination signal exists in a sequence context which can be folded into a hairpin structure in which the entire termination signal is base paired with a complementary sequence in the stem of the hairpin. It seems likely that formation of the hairpin structure inhibits the activity of the termination sequence. Internal early transcription termination signals have been observed in several other early poxvirus genes, including the cowpox virus 132-kDa RNA polymerase subunit (25), the

vaccinia virus DNA polymerase gene (10), the vaccinia virus orf J3 R (27), and the vaccinia virus orf I8 R (Z. Fathi and R. C. Condit, submitted for publication). Only in the case of orf I8 R does the sequence seem to function *in vivo* under normal conditions of infection. Each of these terminators can be folded into secondary structures, though none of the hypothetical structures is as stable as that which is formed in the abortive late gene. These observations raise the possibility that internal early termination signals are used to regulate expression of early poxvirus genes and that the regulation is modulated in part by RNA secondary structure.

Another interesting feature of the early transcript of the abortive late gene is the existence of a 426-nucleotide 5' untranslated region containing only one AUG codon, which would initiate translation of a peptide 2 amino acids in length. Long 5' untranslated regions are not unprecedented in poxvirus mRNAs. The early transcripts of orf I8 R (Fathi and Condit, submitted) and of the 22-kDa RNA polymerase subunit (3) contain 518- and 197-nucleotide 5' untranslated regions, respectively. Translation of these mRNAs is not easily accommodated by the leaky scanning model for translation initiation, in which ribosomes recognize the capped 5' end of an mRNA and scan downstream for the first available AUG in a favored sequence context (15). However, recent evidence suggests that translation of picornavirus mRNA, which contains a long, uncapped 5' untranslated region, may be initiated via RNA sequence-mediated, 5' end-independent ribosome binding (14, 26). It seems possible that translation of vaccinia virus mRNA containing long 5' untranslated regions may also occur by a mechanism other than leaky scanning.

Transcription does not terminate at discrete sites late during vaccinia virus infection, resulting in transcripts with heterogeneous 3' ends (7, 20). In addition, some late poxvirus transcripts have been shown to contain poly(A) at their 5' ends (2, 32). The poly(A) heads are not encoded in the DNA template but may be polymerized during transcription initiation. Synthesis of the poly(A) head may be influenced by the sequence TAAATG (in which the ATG is the translation initiation codon) which is found upstream of many vaccinia virus coding sequences (8, 13, 13a, 16, 29). The failure to terminate late transcription results in readthrough of downstream genes and in the formation of complex double-stranded RNA structures which can potentially cause artifactual signals in both S1 and primer extension analyses (31). We therefore interpret the analyses of the late transcripts reported here with caution. Northern analysis shows clearly that orf Ac R is transcribed both early and late during infection, whereas orfs Aa L and Ab L are transcribed only late during infection. Both S1 and primer extension analyses strongly suggest that the orf Ab L late transcript is initiated immediately upstream of the orf Ab L translation initiation codon and that the transcript has a poly(A) head. Specifically, S1 analysis yields an AUG proximal signal and primer extension yields a ladder of bands diagnostic of a poly(A) head which, according to the data, would begin at the site defined by the S1 analysis. orf Ab L also contains the sequence TAAATG at its 5' end, strengthening the argument that the S1 and primer extension analyses are detecting an authentic poly(A)⁺ 5' mRNA end.

Interpretation of the end analysis of the orf Aa L and orf Ac R late transcripts is less straightforward. Although in both cases some results indicate AUG-proximal mRNA 5' ends, in neither case are the S1 and primer extension results in complete agreement with each other and some of the signals could be artifactual in nature. For example, the

upstream S1 signal for orf Aa L could result from breathing at an AT-rich region of a hybrid formed with an orf Ab L readthrough transcript, and the orf Ac R AUG proximal signals observed in both S1 and primer extension analyses could result from formation of an RNA duplex between orf Ab L and Ac R transcripts. In fact, the data do not exclude the interesting possibilities that the orf Aa L gene product is translated from a bicistronic transcript initiated from orf Ab L or that the orf Ac R late transcript contains a long 5' untranslated region corresponding to the primer extension signals observed with poly(A)⁺ RNA. Precise characterization of the late promoters for these genes must await transcriptional analysis of deleted DNA templates.

ACKNOWLEDGMENTS

We thank Zahra Fathi for the clone pB¹ 14.0 and Ricco Wittek for the clone pBP 3.8. We thank Ed Niles, Zahra Fathi, and Urvashi Hooda-Dhingra for helpful discussions while the work was in progress. We thank Ed Niles for critical reading of the manuscript.

This work was supported by Public Health Service grant AI18094 from the National Institutes of Health.

LITERATURE CITED

1. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. (ed.). 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
2. Bertholet, C., E. Van Meir, B. ten Heggeler-Bordier, and R. Wittek. 1987. Vaccinia virus produces late mRNAs by discontinuous synthesis. *Cell* 50:153-162.
3. Broyles, S. S., and B. Moss. 1986. Homology between RNA polymerases of poxviruses, prokaryotes, and eukaryotes: nucleotide sequence and transcriptional analysis of vaccinia virus genes encoding 147-kDa and 22-kDa subunits. *Proc. Natl. Acad. Sci. USA* 83:3141-3145.
4. Cohrs, R. J., R. C. Condit, R. F. Pacha, C. L. Thompson, and O. K. Sharma. 1989. Modulation of ppp(A2')_nA-dependent RNase by a temperature-sensitive mutant of vaccinia virus. *J. Virol.* 63:948-951.
5. Condit, R. C., and A. Motyczka. 1981. Isolation and preliminary characterization of temperature-sensitive mutants of vaccinia virus. *Virology* 113:224-241.
6. Condit, R. C., A. Motyczka, and G. Spizz. 1983. Isolation, characterization, and physical mapping of temperature-sensitive mutants of vaccinia virus. *Virology* 128:429-443.
7. Cooper, J. A., R. Wittek, and B. Moss. 1981. Extension of the transcriptional and translational map of the left end of the vaccinia virus genome to 21 kilobase pairs. *J. Virol.* 39:733-745.
8. de Magistris, L., and H. G. Stunnenberg. 1988. cis-acting sequences affecting the length of the poly(A) head of vaccinia virus late transcripts. *Nucleic Acids Res.* 16:3141-3156.
9. Devereux, J., P. Haerberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-395.
10. Earl, P. L., E. V. Jones, and B. Moss. 1986. Homology between DNA polymerases of poxviruses, herpesviruses, and adenoviruses: nucleotide sequence of the vaccinia virus DNA polymerase gene. *Proc. Natl. Acad. Sci. USA* 83:3659-3663.
11. Goswami, B. B., and K. Sharma. 1984. Degradation of rRNA in interferon-treated vaccinia virus-infected cells. *J. Biol. Chem.* 259:1371-1374.
12. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.
13. Hirt, P., G. Hiller, and R. Wittek. 1986. Localization and fine structure of a vaccinia virus gene encoding an envelope antigen. *J. Virol.* 58:757-764.
- 13a. Ink, B. S., and D. J. Pickup. 1990. Vaccinia virus directs the synthesis of early messenger RNAs containing 5' poly(A) sequences. *Proc. Natl. Acad. Sci. USA* 87:1536-1540.
14. Jang, S. K., M. V. Davies, R. J. Kaufman, and E. Wimmer.

1989. Initiation of protein synthesis by internal entry of ribosomes into the 5' nontranslated region of encephalomyocarditis virus RNA in vivo. *J. Virol.* **63**:1651-1660.
15. Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* **108**:229-241.
16. Lee-Chen, G. J., N. Bourgeois, K. Davidson, R. Condit, and E. G. Niles. 1988. Structure of the transcription initiation and termination sequences of eight early genes from the vaccinia virus HindIII D fragment. *Virology* **163**:64-79.
17. Lee-Chen, G. J., and E. G. Niles. 1988. Map positions of the 5' ends of eight mRNAs synthesized from the late genes in the vaccinia virus HindIII D fragment. *Virology* **163**:80-92.
18. Lengyl, P. 1982. Biochemistry of interferons and their actions. *Annu. Rev. Biochem.* **51**:251-282.
19. Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. *Science* **227**:1435-1441.
20. Mahr, A., and B. E. Roberts. 1984. Arrangement of late RNAs transcribed from a 7.1-kilobase *EcoRI* vaccinia virus DNA fragment. *J. Virol.* **49**:510-520.
21. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499-580.
22. Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* **101**:20-77.
23. Moss, B. 1985. Replication of poxviruses, p. 685-704. In B. N. Fields, D. M. Knipe, J. L. Melnick, R. M. Chanock, B. R. Roizman, and R. E. Shope (ed.), *Virology*. Raven Press, New York.
24. Pacha, R. F., and R. C. Condit. 1985. Characterization of a temperature-sensitive mutant of vaccinia virus reveals a novel function that prevents virus-induced breakdown of RNA. *J. Virol.* **56**:395-403.
25. Patel, D. D., and D. J. Pickup. 1989. The second-largest subunit of the poxvirus RNA polymerase is similar to the corresponding subunits of procaryotic and eucaryotic RNA polymerases. *J. Virol.* **63**:1076-1086.
26. Pelletier, J., and N. Sonenberg. 1988. Internal initiation of translation of eucaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature (London)* **334**:320-325.
27. Plucienniczak, A., E. Schroeder, G. Zettlmeissl, and R. E. Streeck. 1985. Nucleotide sequence of a cluster of early and late genes in a conserved segment of the vaccinia virus genome. *Nucleic Acids Res.* **13**:985-998.
28. Rohrmann, G., L. Yeun, and B. Moss. 1986. Transcription of vaccinia virus early genes by enzymes isolated from vaccinia virions terminates downstream of a regulatory sequence. *Cell* **46**:1029-1035.
29. Rosel, J. L., P. L. Earl, J. P. Weir, and B. Moss. 1986. Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the HindIII H genome fragment. *J. Virol.* **60**:436-449.
30. Sanger, F., F. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
31. Schmitt, J. F. C., and H. G. Stunnenberg. 1988. Sequence and transcriptional analysis of the vaccinia virus HindIII I fragment. *J. Virol.* **62**:1889-1897.
32. Schwer, B., P. Visca, J. C. Vos, and H. G. Stunnenberg. 1987. Discontinuous transcription or RNA processing of vaccinia virus late messengers results in a 5' poly(A) leader. *Cell* **50**:163-169.
33. Shuman, S., and B. Moss. 1988. Factor-dependent transcription termination by vaccinia virus RNA polymerase. Evidence that the cis-acting termination signal is in nascent RNA. *J. Biol. Chem.* **263**:6220-6225.
34. Staden, R. 1982. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucleic Acids Res.* **10**:4731-4751.
35. Thompson, C. L., and R. C. Condit. 1986. Marker rescue mapping of vaccinia virus temperature-sensitive mutants using overlapping cosmid clones representing the entire virus genome. *Virology* **150**:10-20.
36. Wittek, R., B. Richner, and G. Hiller. 1984. Mapping of the genes coding for the two major vaccinia virus core polypeptides. *Nucleic Acids Res.* **12**:4835-4848.
37. Yuen, L., and B. Moss. 1987. Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes. *Proc. Natl. Acad. Sci. USA* **84**:6417-6421.